09/486882
514 R PCT/PTO 02 MAR 2000

9/PRTS

Chimeric binding peptide library screening method

The present invention relates generally to methods for screening nucleotide libraries for sequences that encode peptides of interest.

Isolating an unknown gene which encodes a desired peptide from a recombinant DNA library can be a difficult task. The use of hybridisation probes may facilitate the process, but their use is generally dependent on knowing at least a portion of the sequence of the gene which encodes the protein. When the sequence is not known, DNA libraries can be expressed in an expression vector, and antibodies have been used to screen for plaques or colonies displaying the desired protein antigen. This procedure has been useful in screening small libraries, but rarely occurring sequences which are represented in less than about 1 in 10^5 clones (as is the case with rarely occurring cDNA molecules or synthetic peptides) can be easily missed, making screening libraries larger than 10^6 clones at best laborious and difficult. Methods designed to address the isolation of rarely occurring sequences by screening libraries of 10^6 clones have been developed and include phage display methods and LacI fusion phage

display, discussed in more detail below.

Phage display methods. Members of DNA libraries which are fused to the N-terminal end of filamentous bacteriophage pIII and pVIII coat proteins have been expressed from an expression vector resulting in the display of foreign peptides on the surface of the phage particle with the DNA encoding the fusion protein packaged in the phage particle (Smith G. P., 1985, Science 228: 1315-1317). The expression vector can be the bacteriophage genome itself, or a phagemid vector, into which a bacteriophage coat protein has been cloned. In the latter case, the host bacterium, containing the phagemid vector, must be co-infected with autonomously replicating bacteriophage, termed helper phage, to provide the full complement of proteins necessary to produce mature phage particles. The helper phage normally has a genetic defect in the origin of replication which results in the preferential packaging of the phagemid genome. Expression of the fusion protein following helper phage infection, allows incorporation of both fusion protein and wild type coat protein into the phage particle during assembly. Libraries of fusion proteins incorporated into phage, can then be selected for binding members against targets of interest (ligands). Bound phage can then be allowed to reinfect *Escherichia coli* (*E. coli*) bacteria and then amplified and the selection repeated, resulting in the enrichment of binding members (Parmley, S. F., & Smith, G. P. 1988., Gene 73: 305-318; Barrett R. W. *et al*., 1992, Analytical Biochemistry 204: 357-364 Williamson *et al*., Proc. Natl. Acad. Sci. USA, 90: 4141-4145; Marks *et al*., 1991, J. Mol. Biol. 222: 581-597).

Several publications describe this method. For example,

US Patent No 5,403,484 describes production of a chimeric protein formed from the viral coat protein and the peptide of interest. In this method at least a functional portion of a viral coat protein is required to cause display of the chimeric protein or a processed form thereof on the outer surface of the virus. In addition, US Patent No 5,571,698 describes a method for obtaining a nucleic acid encoding a binding protein, a key component of which comprises preparing a population of amplifiable genetic packages which have a genetically determined outer surface protein, to cause the display of the potential binding domain on the outer surface of the genetic package. The genetic packages are selected from the group consisting of cells, spores and viruses. For example when the genetic package is a bacterial cell, the outer surface transport signal is derived from a bacterial outer surface protein, and when the genetic package is a filamentous bacteriophage, the outer surface transport signal is provided by the gene pIII (minor coat protein) or pVIII (major coat protein) of the filamentous phage.

WO-A-92/01047 and WO-A-92/20791 describe methods for producing multimeric specific binding pairs, by expressing a first polypeptide chain fused to a viral coat protein, such as the gene pIII protein, of a secreted replicable genetic display package (RGDP) which displays a polypeptide at the surface of the package, and expressing a second polypeptide chain of the multimer, and allowing the two chains to come together as part of the RGDP.

<u>LacI fusion plasmid display</u>. This method is based on the DNA binding ability of the lac repressor. Libraries of random peptides are fused to the lacI repressor

protein, normally to the C-terminal end, through expression from a plasmid vector carrying the fusion gene. Linkage of the LacI-peptide fusion to its encoding DNA occurs via the lacO sequences on the plasmid, forming a stable peptide-LacI-peptide complex. These complexes are released from their host bacteria by cell lysis, and peptides of interest isolated by affinity purification on an immobilised target. The plasmids thus isolated can then be reintroduced into *E. coli* by electroporation to amplify the selected population for additional rounds of screening (Cull, M. G. et al. 1992. Proc. Natl. Acad. Sci. U.S.A. 89:1865-1869).

US Patent No 5498530 describes a method for constructing a library of random peptides fused to a DNA binding protein in appropriate host cells and culturing the host cells under conditions suitable for expression of the fusion proteins intra-cellularly, in the cytoplasm of the host cells. This method also teaches that the random peptide is located at the carboxy terminus of the fusion protein and that the fusion protein-DNA complex is released from the host cell by cell lysis. No method is described for the protection of the DNA from degradation once released from the lysed cell. Several DNA binding proteins are claimed but no examples are shown except lacI.

There remains a need for methods of constructing peptide libraries in addition to the methods described above. For instance, the above methods do not permit production of secreted peptides with a free carboxy terminus. The present invention describes an alternative method for isolating peptides of interest from libraries and has significant advantages over the prior art methods.

In general terms, the present invention provides a method for screening a nucleotide library (usually a DNA library) for a nucleotide sequence which encodes a target peptide of interest. The method involves physically linking each peptide to a polynucleotide including the specific nucleotide sequence encoding that peptide. Linkage of a peptide to its encoding nucleotide sequence is achieved via linkage of the peptide to a nucleotide binding domain. A bifunctional chimeric protein with a nucleotide binding domain and a library member or target peptide (preferably with a function of interest) is thus obtained. The peptide of interest is bound to the polynucleotide encoding that peptide via the nucleotide binding domain of the chimeric protein.

The polynucleotide-chimeric protein complex is then incorporated within a peptide display carrier package (PDCP), protecting the polynucleotide from subsequent degradation, while displaying the target peptide portion on the outer surface of the peptide display carrier package (PDCP).

Thus, in one aspect, the present invention provides a peptide display carrier package (PDCP), said package comprising a polynucleotide-chimeric protein complex wherein the chimeric protein has a nucleotide binding portion and a target peptide portion, wherein said polynucleotide comprises a nucleotide sequence motif which is specifically bound by said nucleotide binding portion, and wherein at least the chimeric protein encoding portion of the polynucleotide not bound by the nucleotide binding portion of the chimeric protein is protected.

In one embodiment the polynucleotide is protected by a

protein which binds non-specifically to naked polynucleotide. Examples include viral coat proteins, many of which are well-known in the art. Where the chosen viral coat protein requires an initiation sequence to commence general binding to the polynucleotide, this will be provided on the polynucleotide at appropriate location(s). A preferred coat protein is coat protein from a bacteriophage, especially M13.

Generally, the nucleic binding portion of the chimeric protein is selected for its specificity for the nucleotide sequence motif present in the recombinant polynucleotide encoding the chimeric protein.

Optionally, the nucleotide sequence motif may be an integral part of the protein encoding region of the polynucleotide. Alternatively, and more usually, the motif may be present in a non-coding region of the polynucleotide. For the purposes of this invention, all that is required is for the motif to be located on the polynucleotide such that the nucleotide binding portion of the chimeric protein is able to recognise and bind to it. Desirably the polynucleotide-chimeric protein complex has a dissociation constant of at least one hour.

Optionally, the recombinant polynucleotide may comprise two or more nucleotide sequence motifs, each of which will be bound by a chimeric protein molecule. Preferably, the motifs are positioned along the length of the polynucleotide to avoid steric hindrance between the bound chimeric proteins.

Preferably, the nucleotide sequence motif is not affected by the presence of additional nucleotide

sequence (e.g. encoding sequence) at its 5' and/or 3' ends. Thus the chimeric fusion protein may include a target peptide portion at its N terminal end, at its C terminal end or may include two target peptide portions (which may be the same or different) at each end of the nucleotide binding portion, ie at both the N and C terminal ends of the chimeric protein. For example one target peptide may be an antibody of known specificity and the other peptide may be a peptide of potential interest.

Desirably the target peptide portion of the chimeric protein is displayed externally on the peptide display carrier package, and is thus available for detection, reaction and/or binding.

In more detail the PDCP may be composed two distinct elements:

a. A polynucleotide-chimeric protein complex. This links the displayed target peptide portion to the polynucleotide encoding that peptide portion through a specific polynucleotide binding portion. The nucleotide sequence encoding the chimeric protein, and the specific nucleotide sequence motif recognised by the nucleotide binding portion of the chimeric protein must be present on a segment of polynucleotide which can be incorporated into the PDCP; and

b. A protective coat. This may be supplied by a replicable carrier or helper package capable of independent existence. Alternatively, a coat protein could be encoded by the recombinant polynucleotide of the invention. The protective coat for the polynucleotide-chimeric protein complex may be composed of a biological material such as protein or lipid, but the protective coat

is not required for linking the target peptide to the polynucleotide encoding that peptide. The protective coat must allow the display of the target peptide portion of the chimeric protein on its outer surface. The carrier or helper package may also provide the mechanism for releasing the intact PDCP from host cells when so required. By way of example, when a bacteriophage is the replicable carrier package, a protein coat of the bacteriophage surrounds the polynucleotide-chimeric protein complex to form the PDCP, which is then extruded from the host bacterial cell.

The invention described herein demonstrates that peptides fused to a nucleotide binding domain can be displayed externally, even through a bacteriophage carrier package protein coat, while still bound to the polynucleotide encoding the displayed peptide.

The present invention also provides a recombinant polynucleotide comprising a nucleotide sequence encoding a chimeric protein having a nucleotide binding portion operably linked to a target peptide portion, wherein said polynucleotide includes a specific nucleotide sequence motif which is bound by the nucleotide binding portion of said chimeric protein and further encoding a non-sequence-specific nucleotide binding protein.

Desirably, the recombinant polynucleotide is a recombinant expression system, able to express the chimeric protein when placed in a suitable environment, for example a compatible host cell. After its expression, the chimeric protein binds to the specific nucleotide sequence (motif) present in the polynucleotide comprising the nucleotide sequence

encoding the chimeric protein.

Optionally there may be a linker sequence located between the nucleotide sequence encoding the nucleotide binding portion and the polynucleotide inserted into the restriction enzyme site of the construct.

Desirably the nucleotide binding portion is a DNA binding domain of an oestrogen or progesterone receptor, or a functional equivalent thereof. Examples of sequences encoding such nucleotide binding portions are set out in SEQ ID Nos 11 and 13.

The term "expression system" is used herein to refer to a genetic sequence which includes a protein-encoding region and is operably linked to all of the genetic signals necessary to achieve expression of that region. Optionally, the expression system may also include regulatory elements, such as a promoter or enhancer to increase transcription and/or translation of the protein encoding region or to provide control over expression. The regulatory elements may be located upstream or downstream of the protein encoding region or within the protein encoding region itself. Where two or more distinct protein encoding regions are present these may use common regulatory element(s) or have separate regulatory element(s).

Generally, the recombinant polynucleotide described above will be DNA. Where the expression system is based upon an M13 vector, usually the polynucleotide binding portion of the expressed chimeric portion will be single-stranded DNA. However, other vector systems may be used and the nucleotide binding portion may be selected to bind preferentially to double-stranded DNA or to double or single-stranded RNA, as convenient.

Additionally the present invention provides a vector containing such a recombinant expression system and host cells transformed with such a recombinant expression system (optionally in the form of a vector).

Whilst the recombinant polynucleotide described above forms an important part of the present invention, we are also concerned with the ability to screen large (e.g. of at least 10^5 members, for example 10^6 or even 10^7 members) libraries of genetic material. One of the prime considerations therefore is the provision of a recombinant genetic construct into which each member of said library can individually be incorporated to form the recombinant polynucleotide described above and to express the chimeric protein thereby encoded (the target peptide of which is encoded by the nucleotide library member incorporated into the construct).

Thus viewed in a further aspect the present invention provides a genetic construct or set of genetic constructs comprising a polynucleotide having a sequence which includes:

i) a sequence encoding a nucleotide binding portion able to recognise and bind to a specific sequence motif;
ii) the sequence motif recognised and bound by the nucleotide binding portion encoded by (i);
iii) a restriction enzyme site which permits insertion of a polynucleotide, said site being designed to operably link said polynucleotide to the sequence encoding the nucleotide binding portion so that expression of the operably linked polynucleotide sequences yields a chimeric protein; and
iv) a sequence encoding a nucleotide binding protein which binds non-specifically to naked

polynucleotide.

Optionally there may be a linker sequence located between the nucleotide sequence encoding the nucleotide binding portion and the sequence of the polynucleotide from the library inserted into the restriction enzyme site of the construct.

Desirably the nucleotide binding portion is a DNA binding domain of an oestrogen or progesterone receptor, or a functional equivalent thereof. Examples of sequences encoding such nucleotide binding portions are set out in SEQ ID Nos 11 and 13.

Suitable genetic constructs according to the invention include pDM12, pDM14 and pDM16, deposited at NCIMB on 28 August 1998 under Nos NCIMB 40970, NCIMB 40971 and NCIMB 40972 respectively.

It is envisaged that a conventionally produced genetic library may be exposed to the genetic construct(s) described above. Thus, each individual member of the genetic library will be separately incorporated into the genetic construct and the library will be present in the form of a library of recombinant polynucleotides (as described above), usually in the form of vectors, each recombinant polynucleotide including as library member.

Thus, in a further aspect, the present invention provides a library of recombinant polynucleotides (as defined above) wherein each polynucleotide includes a polynucleotide obtained from a genetic library and which encodes the target peptide portion of the chimeric protein expressed by the recombinant polynucleotide.

Optionally, the chimeric protein may further include a linker sequence located between the nucleotide binding portion and the target peptide portion. The linker sequence will reduce steric interference between the two portions of the protein. Desirably the linker sequence exhibits a degree of flexibility.

Also disclosed are methods for constructing and screening libraries of PDCP particles, displaying many different peptides, allowing the isolation and identification of particular peptides by means of affinity techniques relying on the binding activity of the peptide of interest. The resulting polynucleotide sequences can therefore be more readily identified, re-cloned and expressed.

A method of constructing a genetic library, said method comprising:

a) constructing multiple copies of a recombinant vector comprising a polynucleotide sequence which encodes a nucleotide binding portion able to recognise and bind to a specific sequence motif (and optionally also including the specific sequence motif);

b) operably linking each said vector to a polynucleotide encoding a target polypeptide, such that expression of said operably linked vector results in expression of a chimeric protein comprising said target peptide and said nucleotide binding portions; wherein said multiple copies of said operably linked vectors collectively express a library of target peptide portions;

c) transforming host cells with the vectors of step

b);

d) culturing the host cells of step c) under conditions suitable for expression of said chimeric protein;

e) providing a recombinant polynucleotide comprising the nucleotide sequence motif specifically recognised by the nucleotide binding portion and exposing this polynucleotide to the chimeric protein of step d) to yield a polynucleotide-chimeric protein complex; and

f) causing production of a non-sequence-specific moiety able to bind to the non-protected portion of the polynucleotide encoding the chimeric protein to form a peptide display carrier package.

The present invention further provides a method of screening a genetic library, said method comprising:

a) exposing the polynucleotide members of said library to multiple copies of a genetic construct comprising a nucleotide sequence encoding a nucleotide binding portion able to recognise and bind to a specific sequence motif, under conditions suitable for the polynucleotides of said library each to be individually ligated into one copy of said genetic construct, to create a library of recombinant polynucleotides;

b) exposing said recombinant polynucleotides to a population of host cells, under conditions suitable for transformation of said host cells by said recombinant polynucleotides;

c) selecting for transformed host cells;

d) exposing said transformed host cells to conditions suitable for expression of said recombinant polynucleotide to yield a chimeric protein; and

e) providing a recombinant polynucleotide comprising the nucleotide sequence motif specifically recognised by the nucleotide binding portion and exposing this polynucleotide to the chimeric protein of step d) to yield a polynucleotide-chimeric protein complex;

f) protecting any exposed portions of the polynucleotide in the complex of step e) to form a peptide display carrier package; and

g) screening said peptide display carrier package to select only those packages displaying a target peptide portion having the characteristics required.

Desirably in step a) the genetic construct is pDM12, pDM14 or pDM16.

Desirably in step f) the peptide display package carrier is extruded from the transformed host cell without lysis of the host cell.

Generally the transformed host cells will be plated out or otherwise divided into single colonies following transformation and prior to expression of the chimeric protein.

The screening step g) described above may look for a particular target peptide either on the basis of

function (e.g. enzymic activity) or structure (e.g. binding to a specific antibody). Once the peptide display carrier package is observed to include a target peptide with the desired characteristics, the polynucleotide portion thereof (which of course encodes the chimeric protein itself) can be amplified, cloned and otherwise manipulated using standard genetic engineering techniques.

The current invention differs from the prior art teaching of the previous disclosures US Patent No 5,403,484 and US Patent No 5,571,698, as the invention does not require outer surface transport signals, or functional portions of viral coat proteins, to enable the display of chimeric binding proteins on the outer surface of the viral particle or genetic package.

The current invention also differs from the teaching of WO-A-92/01047 and WO-A-92/20791, as no component of a secreted replicable genetic display package, or viral coat protein is required, to enable display of the target peptide on the outer surface of the viral particle.

The current invention differs from the teaching of US Patent No 5498530, as it enables the display of chimeric proteins, linked to the polynucleotide encoding the chimeric protein, extra-cellularly, not in the cytoplasm of a host cell. In the current invention the chimeric proteins are presented on the outer surface of a peptide display carrier package (PDCP) which protects the DNA encoding the chimeric protein, and does not require cell lysis to obtain access to the chimeric protein-DNA complex. Finally, the current invention does not rely upon the lacI DNA binding protein to form the chimeric protein-DNA complex.

In one embodiment of the invention, the nucleotide binding portion of the chimeric protein comprises a DNA binding domain from one or more of the nuclear steroid receptor family of proteins, or a functional equivalent of such a domain. Particular examples include (but are not limited to) a DNA binding domain of the oestrogen receptor or the progesterone receptor, or functional equivalents thereof. These domains can recognise specific DNA sequences, termed hormone response elements (HRE), which can be bound as both double and single-stranded DNA. The DNA binding domain of such nuclear steroid receptor proteins is preferred.

The oestrogen receptor is especially referred to below by way of example, for convenience since:
(a) The oestrogen receptor is a large multifunctional polypeptide of 595 amino acids which functions in the cytoplasm and nucleus of eukaryotic cells (Green et al., 1986, Science 231: 1150-1154). A minimal high affinity DNA binding domain (DBD) has been defined between amino acids 176 and 282 (Mader et al., 1993, Nucleic Acids Res. 21: 1125-1132). The functioning of this domain (i.e. DNA binding) is not inhibited by the presence of non-DNA binding domains at both the N and C terminal ends of this domain, in the full length protein.

(b) The oestrogen receptor DNA binding domain fragment (amino acids 176-282) has been expressed in *E. coli* and shown to bind to the specific double stranded DNA oestrogen receptor target HRE nucleotide sequence, as a dimer with a similar affinity (0.5nM) to the parent molecule (Murdoch et al. 1990, Biochemistry 29: 8377-8385; Mader et al., 1993, Nucleic Acids Research 21: 1125-1132). DBD dimerization on the surface of the PDCP should result in two peptides displayed per particle.

This bivalent display can aid in the isolation of low affinity peptides and peptides that are required to form a bivalent conformation in order to bind to a particular target, or activate a target receptor. The oestrogen receptor is capable of binding to its 38 base pair target HRE sequence, consensus sequence:

1) 5'-TCAGGTCAGAGTGACCTGAGCTAAAATAACACATTCAG-3'
("minus strand") SEQ ID No 1, and

2) 3'-AGTCCAGTCTCACTGGACTCGATTTTATTGTGTAAGTC-5'
("plus strand") SEQ ID No 2,

with high affinity and specificity, under the salt and pH conditions normally required for selection of binding peptides. Moreover, binding affinity is increased 60-fold for the single-stranded coding, or "plus", strand (i.e. SEQ ID No 2) of the HRE nucleotide sequence over the double stranded form of the specific target nucleotide sequence (Peale et al. 1988, Proc. Natl. Acad. Sci. USA 85: 1038-1042; Lannigan & Notides, 1989, Proc. Natl. Acad. Sci. USA 86: 863-867).

In an embodiment of the invention where the DNA binding component of the peptide display carrier package is the oestrogen receptor, the nucleotide (DNA) binding portion contains a minimum sequence of amino acids 176-282 of the oestrogen receptor protein. In addition, the consensus oestrogen receptor target HRE sequence is cloned in such a way that if single stranded DNA can be produced then the coding, or "plus", strand of the oestrogen receptor HRE nucleotide sequence is incorporated into single-stranded DNA. An example of a vector suitable for this purpose is pUC119 (see Viera et al., Methods in Enzymology, Vol 153, pages 3-11, 1987).

In a preferred embodiment of the invention a peptide display carrier package (PDCP) can be assembled when a bacterial host cell is transformed with a bacteriophage vector, which vector comprises a recombinant polynucleotide as described above. The expression vector will also comprise the specific nucleotide motif that can be bound by the nucleotide binding portion of the chimeric protein. Expression of recombinant polynucleotide results in the production of the chimeric protein which comprises the target peptide and the nucleotide binding portion. The host cell is grown under conditions suitable for chimeric protein expression and assembly of the bacteriophage particles, and the association of the chimeric protein with the specific nucleotide sequence in the expression vector.

In this embodiment, since the vector is a bacteriophage, which replicates to produce a single-stranded DNA, the nucleotide binding portion preferably has an affinity for single-stranded DNA. Incorporation of the vector single-stranded DNA-chimeric protein complex into bacteriophage particles results in the assembly of the peptide display carrier package (PDCP), and display of the target peptide on the outer surface of the PDCP.

In this embodiment both of the required elements for producing peptide display carrier packages are contained on the same vector. Incorporation of the DNA-chimeric protein complex into a peptide display carrier package (PDCP) is preferred as DNA degradation is prevented, large numbers of PDCPs are produced per host cell, and the PDCPs are easily separated from the host cell without recourse to cell lysis.

In a more preferred embodiment, the vector of the is a

phagemid vector (for example pUC119) where expression of the chimeric protein is controlled by an inducible promoter. In this embodiment the PDCP can only be assembled following infection of the host cell with both phagemid vector and helper phage. The transfected host cell is then cultivated under conditions suitable for chimeric protein expression and assembly of the bacteriophage particles.

In this embodiment the elements of the PDCP are provided by two separate vectors. The phagemid derived PDCP is superior to phagemid derived display packages disclosed in WO-A-92/01047 where a proportion of packages displaying bacteriophage coat protein fusion proteins will contain the helper phage DNA, not the fusion protein DNA sequence. In the current invention, a PDCP can display the chimeric fusion protein only when the package contains the specific nucleotide motif recognised by the nucleotide binding portion. In most embodiments this sequence will be present on the same DNA segment that encodes the fusion protein. In addition, the prior art acknowledges that when mutant and wild type proteins are co-expressed in the same bacterial cell, the wild type protein is produced preferentially. Thus, when the wild type helper phage, phage display system of WO-A-92/01047 is used, both wild type gene pIII and target peptide-gene pIII chimeric proteins are produced in the same cell. The result of this is that the wild type gene pIII protein is preferentially packaged into bacteriophage particles, over the chimeric protein. In the current invention, there is no competition with wild type bacteriophage coat proteins for packaging.

Desirably the target peptide is displayed in a location exposed to the external environment of the PDCP, after

the PDCP particle has been released from the host cell without recourse to cell lysis. The target peptide is then accessible for binding to its ligand. Thus, the target peptide may be located at or near the N-terminus or the C-terminus of a nucleotide binding domain, for example the DNA binding domain of the oestrogen receptor.

The present invention also provides a method for screening a DNA library expressing one or more polypeptide chains that are processed, folded and assembled in the periplasmic space to achieve biological activity. The PDCP may be assembled by the following steps:

(a) Construction of N- or C-terminal DBD chimeric protein fusions in a phagemid vector.

(i) When the target peptide is located at the N-terminus of the nucleotide binding portion, a library of DNA sequences each encoding a potential target peptide is cloned into an appropriate location of an expression vector (i.e. behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of the downstream fusion protein to the periplasmic space) and upstream of the sequence encoding the nucleotide binding portion. In a preferred embodiment the DNA sequence(s) of interest may be joined, by a region of DNA encoding a flexible amino acid linker, to the 5'-end of an oestrogen receptor DBD.

(ii) Alternatively, when the target peptide is located at the C-terminus of the nucleotide binding domain, a library of DNA sequences each encoding a potential target peptide is cloned into the expression vector so that the nucleotide sequence coding for the nucleotide binding portion is upstream of the cloned

DNA target peptide encoding sequences, said nucleotide binding portion being positioned behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of the downstream fusion protein to the periplasmic space. In a preferred embodiment, DNA sequence(s) of interest may be joined, by a region of DNA encoding a flexible amino acid linker oestrogen receptor DBD DNA sequence.

Located on the expression vector is the specific HRE nucleotide sequence recognised, and bound, by the oestrogen receptor DBD. In order to vary the number of chimeric proteins displayed on each PDCP particle, this sequence can be present as one or more copies in the vector.

(b) Incorporation into the PDCP. Non-lytic helper bacteriophage infects host cells containing the expression vector. Preferred types of bacteriophage include the filamentous phage fd, f1 and M13. In a more preferred embodiment the bacteriophage may be M13K07.

The protein(s) of interest are expressed and transported to the periplasmic space, and the properly assembled proteins are incorporated into the PDCP particle by virtue of the high affinity interaction of the DBD with the specific target nucleotide sequence present on the phagemid vector DNA which is naturally packaged into phage particles in a single-stranded form. The high affinity interaction between the DBD protein and its specific target nucleotide sequence prevents displacement by bacteriophage coat proteins resulting in the incorporation of the protein(s) of interest onto the surface of the PDCP as it is extruded from the cell.

(c) Selection of the peptide of interest. Particles which display the peptide of interest are then selected from the culture by affinity enrichment techniques. This is accomplished by means of a ligand specific for the protein of interest, such as an antigen if the protein of interest is an antibody. The ligand may be presented on a solid surface such as the surface of an ELISA plate, or in solution. Repeating the affinity selection procedure provides an enrichment of clones encoding the desired sequences, which may then be isolated for sequencing, further cloning and/or expression.

Numerous types of libraries of peptides fused to the DBD can be screened under this embodiment including:

(i) Random peptide sequences encoded by synthetic DNA of variable length.

(ii) Single-chain Fv antibody fragments. These consist of the antibody heavy and light chain variable region domains joined by a flexible linker peptide to create a single-chain antigen binding molecule.

(iii) Random fragments of naturally occurring proteins isolated from a cell population containing an activity of interest.

In another embodiment the invention concerns methods for screening a DNA library whose members require more than one chain for activity, as required by, for example, antibody Fab fragments for ligand binding. In this embodiment heavy or light chain antibody DNA is joined to a nucleotide sequence encoding a DNA binding domain of, for example, the oestrogen receptor in a

phagemid vector. Typically the antibody DNA library sequences for either the heavy (VH and CH1) or light chain (VL and CL) genes are inserted in the 5' region of the oestrogen receptor DBD DNA, behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of the downstream fusion protein to the periplasmic space.

Thus, a DBD fused to a DNA library member-encoded protein is produced and assembled in to the viral particle after infection with bacteriophage. The second and any subsequent chain(s) are expressed separately either:

(a) from the same phagemid vector containing the DBD and the first polypeptide fusion protein,
or

(b) from a separate region of DNA which may be present in the host cell nucleus, or on a plasmid, phagemid or bacteriophage expression vector that can co-exist, in the same host cell, with the first expression vector, so as to be transported to the periplasm where they assemble with the first chain that is fused to the DBD protein as it exits the cell. Peptide display carrier packages (PDCP) which encode the protein of interest can then be selected by means of a ligand specific for the protein.

In yet another embodiment, the invention concerns screening libraries of bi-functional peptide display carrier packages where two or more activities of interest are displayed on each PDCP. In this embodiment, a first DNA library sequence(s) is inserted next to a first DNA binding domain (DBD) DNA sequence,

for example the oestrogen receptor DBD, in an appropriate vector, behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of this first chimeric protein to the periplasmic space. A second chimeric protein is also produced from the same, or separate, vector by inserting a second DNA library sequence(s) next to a second DBD DNA sequence which is different from the first DBD DNA sequence, for example the progesterone receptor DBD, behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader. The first, or only, vector contains the specific HRE nucleotide sequences for both oestrogen and progesterone receptors. Expression of the two chimeric proteins, results in a PDCP with two different chimeric proteins displayed. As an example, one chimeric protein could possess a binding activity for a particular ligand of interest, while the second chimeric protein could possess an enzymatic activity. Binding by the PDCP to the ligand of the first chimeric protein could then be detected by subsequent incubation with an appropriate substrate for the second chimeric protein. In an alternative embodiment a bi-functional PDCP may be created using a single DBD, by cloning one peptide at the 5'-end of the DBD, and a second peptide at the 3'-end of the DBD. Expression of this single bi-functional chimeric protein results in a PDCP with two different activities.

We have investigated the possibility of screening libraries of peptides, fused to a DNA binding domain and displayed on the surface of a display package, for particular peptides with a biological activity of interest and recovering the DNA encoding that activity. Surprisingly, by manipulating the oestrogen receptor

DNA binding domain in conjunction with M13 bacteriophage we have been able to construct novel particles which display large biologically functional molecules, that allows enrichment of particles with the desired specificity.

The invention described herein provides a significant breakthrough in DNA library screening technology.

The invention will now be further described by reference to the non-limiting examples and figures below.

Description of Figures

Figure 1 shows the pDM12 N-terminal fusion oestrogen receptor DNA binding domain expression vector nucleotide sequence (SEQ ID No 3), between the HindIII and EcoRI restriction sites, comprising a pelB leader secretion sequence (in italics), multiple cloning site containing SfiI and NotI sites, flexible $(glycine)_4$-serine linker sequence (boxed), a fragment of the oestrogen receptor gene comprising amino acids 176-282 (SEQ ID No 4) of the full length molecule, and the 38 base pair consensus oestrogen receptor DNA binding domain HRE sequence.

Figure 2 shows the OD_{450nm} ELISA data for negative control M13K07 phage, and single-clone PDCP display culture supernatants (#1-4, see Example 3) isolated by selection of the lymphocyte cDNA-pDM12 library against anti-human immunoglobulin kappa antibody.

Figure 3 shows partial DNA (SEQ ID No 5) and amino acid (SEQ ID No 6) sequence for the human immunoglobulin kappa constant region (Kabat, E. A. et al., Sequences

of Proteins of Immunological Interest. 4th edition. U.S. Department of Health and Human Services. 1987), and ELISA positive clones #2 (SEQ ID Nos 7 and 8) and #3 (SEQ ID Nos 9 and 10) from Figure 2 which confirms the presence of human kappa constant region DNA in-frame with the pelB leader sequence (pelB leader sequence is underlined, the leader sequence cleavage site is indicated by an arrow). The differences in the 5'-end sequence demonstrates that these two clones were selected independently from the library stock. The PCR primer sequence is indicated in bold, clone #2 was originally amplified with CDNAPCRBAK1 and clone #3 was amplified with CDNAPCRBAK2.

Figure 4 shows the pDM14 N-terminal fusion oestrogen receptor DNA binding domain expression vector nucleotide sequence (SEQ ID No 11), between the HindIII and EcoRI restriction sites, comprising a pelB leader secretion sequence (in italics), multiple cloning site containing SfiI and NotI sites, flexible (glycine)$_4$-serine linker sequence (boxed), a fragment of the oestrogen receptor gene comprising amino acids 176-282 (SEQ ID No 12) of the full length molecule, and the two 38 base pair oestrogen receptor DNA binding domain HRE sequences (HRE 1 and HRE 2).

Figure 5 shows the pDM16 C-terminal fusion oestrogen receptor DNA binding domain expression vector nucleotide sequence (SEQ ID No 13), between the HindIII and EcoRI restriction sites, comprising a pelB leader secretion sequence (in italics), a fragment of the oestrogen receptor gene comprising amino acids 176-282 (SEQ ID No 14) of the full length molecule, flexible (glycine)$_4$-serine linker sequence (boxed), multiple cloning site containing SfiI and NotI sites and the 38 base pair oestrogen receptor DNA binding domain HRE

sequence.

Figure 6 shows the OD_{450nm} ELISA data for N-cadherin-pDM16 C-terminal display PDCP binding to anti-pan-cadherin monoclonal antibody in serial dilution ELISA as ampicillin resitance units (a.r.u.). Background binding of negative control M13K07 helper phage is also shown.

Figure 7 shows the OD_{450nm} ELISA data for *in vivo* biotinylated PCC-pDM16 C-terminal display PDCP binding to streptavidin in serial dilution ELISA as ampicillin resitance units (a.r.u.). Background binding of negative control M13K07 helper phage is also shown.

Figure 8 shows the OD_{450nm} ELISA data for a human scFv PDCP isolated from a human scFv PDCP display library selected against substance P. The PDCP was tested against streptavidin (1), streptavidin-biotinylated substance P (2), and streptavidin-biotinylated CGRP (3), in the presence (B) or absence (A) of free substance P.

Figure 9 shows the DNA (SEQ ID Nos 15 and 17) and amino acid (SEQ ID No 16 and 18) sequence of the substance P binding scFv isolated from a human scFv PDCP display library selected against substance P. Heavy chain (SEQ ID Nos 15 and 16) and light chain (SEQ ID Nos 17 and 18) variable region sequence is shown with the CDRs underlined and highlighted in bold.

Materials and Methods

The following procedures used by the present applicant are described in Sambrook, J., et al., 1989 supra.: restriction enzyme digestion, ligation, preparation of electrocompetent cells, electroporation, analysis of

restriction enzyme digestion products on agarose gels, DNA purification using phenol/chloroform, preparation of 2xTY medium and plates, preparation of ampicillin, kanamycin, IPTG (Isopropyl β-D-Thiogalactopyranoside) stock solutions, and preparation of phosphate buffered saline.

Restriction enzymes, T4 DNA ligase and cDNA synthesis reagents (Superscript plasmid cDNA synthesis kit) were purchased from Life Technologies Ltd (Paisley, Scotland, U.K.). Oligonucleotides were obtained from Cruachem Ltd (Glasgow, Scotland, U.K.), or Genosys Biotechnologies Ltd (Cambridge, U.K.). Taq DNA polymerase, Wizard SV plasmid DNA isolation kits, streptavidin coated magnetic beads and mRNA isolation reagents (PolyATract 1000) were obtained from Promega Ltd (Southampton, Hampshire, U.K.). Taqplus DNA polymerase was obtained from Stratagene Ltd (Cambridge, U.K.). PBS, BSA, streptavidin, substance P and anti-pan cadherin antibody were obtained from SIGMA Ltd (Poole, Dorset, U.K.). Anti-M13-HRP conjugated antibody, Kanamycin resistant M13K07 helper bacteriophage and RNAguard were obtained from Pharmacia Ltd (St. Albans, Herts, U.K.) and anti-human Igκ antibody from Harlan-Seralab (Loughborough, Leicestershire, U.K.) Biotinylated substance P and biotinylated calcitonin gene related peptide (CGRP) were obtained from Peninsula Laboratories (St. Helens, Merseyside, U.K.).

Specific embodiments of the invention are given below in Examples 1 to 9.

Example 1. Construction of a N-terminal PDCP display phagemid vector pDM12.

The pDM12 vector was constructed by inserting an oestrogen receptor DNA binding domain, modified by appropriate PCR primers, into a phagemid vector pDM6. The pDM6 vector is based on the pUC119 derived phage display vector pHEN1 (Hoogenboom et al., 1991, Nucleic Acids Res. 19: 4133-4137). It contains $(Gly)_4Ser$ linker, Factor Xa cleavage site, a full length gene III, and streptavidin tag peptide sequence (Schmidt, T.G. and Skerra, A., 1993, Protein Engineering 6: 109-122), all of which can be removed by NotI-EcoRI digestion and agarose gel electrophoresis, leaving a pelB leader sequence, SfiI, NcoI and PstI restriction sites upstream of the digested NotI site. The cloned DNA binding domain is under the control of the lac promoter found in pUC119.

Preparation of pDM6

The pDM12 vector was constructed by inserting an oestrogen receptor DNA binding domain, modified by appropriate PCR primers, into a phagemid vector pDM6. The pDM6 vector is based on the gene pIII phage display vector pHEN1 (Hoogenboom et al., 1991, Nucleic Acids Res. 19: 4133-4137), itself derived from pUC119 (Viera, J. and Messing, J., 1987, Methods in Enzymol. 153: 3-11). It was constructed by amplifying the pIII gene in pHEN1 with two oligonucleotides:

PDM6BAK: 5 -TTT TCT GCA GTA ATA GGC GGC CGC AGG GGG AGG AGG GTC CAT CGA AGG TCG CGA AGC AGA GAC TGT TGA AAG T-3 (SEQ ID No 19) and

PDM6FOR: 5 - TTT TGA ATT CTT ATT AAC CAC CGA ACT GCG

GGT GAC GCC AAG CGC TTG CGG CCG TTA AGA CTC CTT ATT ACG CAG-3 (SEQ ID No 20).

and cloning the PstI-EcoRI digested PCR product back into similarly digested pHEN1, thereby removing the c-myc tag sequence and supE TAG codon from pHEN1. The pDM6 vector contains a $(Gly)_4$Ser linker, Factor Xa cleavage site, a full length gene III, and streptavidin tag peptide sequence (Schmidt, T.G. and Skerra, A., 1993, Protein Engineering 6: 109-122), all of which can be removed by NotI-EcoRI digestion and agarose gel electrophoresis, leaving a pelB leader sequence, SfiI, NcoI and PstI restriction sites upstream of the digested NotI site. The cloned DNA binding domain is under the control of the lac promoter found in pUC119.

The oestrogen receptor DNA binding domain was isolated from cDNA prepared from human bone marrow (Clontech, Palo Alto, California, U.S.A.). cDNA can be prepared by many procedures well known to those skilled in the art. As an example, the following method using a Superscript plasmid cDNA synthesis kit can be used:

(a) First strand synthesis.

5μg of bone marrow mRNA, in 5μl DEPC-treated water was thawed on ice and 2μl (50pmol) of cDNA synthesis primer (5'-AAAAGCGGCCGCACTGGCCTGAGAGA$(N)_6$-3') (SEQ ID No 21) was added to the mRNA and the mixture heated to 70°C for 10 minutes, then snap-chilled on ice and spun briefly to collect the contents to the bottom of the tube. The following were then added to the tube:

1000u/ml RNAguard	1μl
5x first strand buffer	4μl
0.1M DTT	2μl
10mM dNTPs	1μl

200u/μl SuperScript II reverse transcriptase	5μl

The mixture was mixed by pipetting gently and incubated at 37°C for 1 hour, then placed on ice.

(b) Second strand synthesis.

The following reagents were added to the first strand reaction:

DEPC-treated water	93μl
5x second strand buffer	30μl
10mM dNTPs	3μl
10u/μl *E. coli* DNA ligase	1μl
10u/μl *E. coli* DNA polymerase	4μl
2u/μl *E. coli* RNase H	1μl

The reaction was vortex mixed and incubated at 16°C for 2 hours. 2μl (10u) of T4 DNA polymerase was added and incubation continued at 16°C for 5 minutes. The reaction was placed on ice and 10μl 0.5M EDTA added, then phenol-chloroform extracted, precipitated and vacuum dried.

(c) Sal I adaptor ligation.

The cDNA pellet was resuspended in 25μl DEPC-treated water, and ligation set up as follows.

cDNA	25μl
5x T4 DNA ligase buffer	10μl
1μg/μl *Sal* I adapters*	10μl
1u/μlT4 DNA ligase	5μl

*Sal I adapters: TCGACCCACGCGTCCG-3' (SEQ ID No 22)
GGGTGCCGAGGC-5' (SEQ ID No 23)

The ligation was mixed gently and incubated for 16 hours at 16°C, then phenol-chloroform extracted, precipitated and vacuum dried. The cDNA/adaptor pellet was resuspended in 41μl of DEPC-treated water and digested with 60 units of NotI at 37°C for 2 hours,

then phenol-chloroform extracted, precipitated and vacuum dried. The cDNA pellet was re-dissolved in 100μl TEN buffer (10mM Tris pH 7.5, 0.1mM EDTA, 25mM NaCl) and size fractionated using a Sephacryl S-500 HR column to remove unligated adapters and small cDNA fragments (<400bp) according to the manufacturers instructions. Fractions were checked by agarose gel electrophoresis and fractions containing cDNA less than 400 base pairs discarded, while the remaining fractions were pooled.

(d) PCR amplification of oestrogen receptor DNA binding domain.

The oestrogen receptor was PCR amplified from 5μl (150-250ng) of bone marrow cDNA using 25pmol of each of the primers pDM12FOR (SEQ ID No 24) (5'-AAAAGAATTCTGAATGTGTTATTTTAGCTCAGGTCACTCTGACCTGATTATCAAGACCCCACTTCACCCCCT) and pDM12BAK (SEQ ID No 25) (5'-AAAAGCGGCCGCAGGGGGAGGAGGGTCCATGGAATCTGCCAAGGAG-3') in two 50μl reactions containing 0.1mM dNTPs, 2.5 units Taq DNA polymerase, and 1x PCR reaction buffer (10mM Tris-HCl pH 9.0, 5mM KCl, 0.01% Triton X-100, 1.5mM $MgCl_2$) (Promega Ltd, Southampton, U.K.). The pDM12FOR primer anneals to the 3'-end of the DNA binding domain of the oestrogen receptor and incorporates two stop codons, the 38 base pair consensus oestrogen receptor HRE sequence, and an EcoRI restriction site. The pDM12BAK primer anneals to the 5'-end of the DNA binding domain of the oestrogen receptor and incorporates the $(Gly)_4Ser$ linker and the NotI restriction site.

Reactions were overlaid with mineral oil and PCR carried out on a Techne PHC-3 thermal cycler for 30 cycles of 94°C, 1 minute; 65°C, 1 minute; 72°C, 1 minute. Reaction products were electrophoresed on an

agarose gel, excised and products purified from the gel using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(e) Restriction digestion and ligation.

The PCR reaction appended NotI and EcoRI restriction sites, the $(Gly)_4$Ser linker, stop codons and the 38 base pair oestrogen receptor target HRE nucleotide sequence to the oestrogen receptor DNA binding domain sequence (see Figure 1). The DNA PCR fragment and the target pDM6 vector (approximately 500ng) were NotI and EcoRI digested for 1 hour at 37°C, and DNA purified by agarose gel electrophoresis and extraction with Geneclean II kit (Bio101, La Jolla, California, U.S.A.). The oestrogen receptor DNA binding domain cassette was ligated into the NotI-EcoRI digested pDM6 vector overnight at 16°C, phenol/chloroform extracted and precipitated then electroporated into TG1 *E. coli* (genotype: K12, (Δlac-pro), supE, thi, hsD5/F'traD36, $proA^+B^+$, $LacI^q$, LacZΔ15) and plated onto 2xTY agar plates supplemented with 1% glucose and 100μg/ml ampicillin. Colonies were allowed to grow overnight at 37°C. Individual colonies were picked into 5ml 2xTY supplemented with 1% glucose and 100μg/ml ampicillin and grown overnight at 37°C. Double stranded phagemid DNA was isolated with a Wizard SV plasmid DNA isolation kit and the sequence confirmed with a Prism dyedeoxy cycle sequencing kit (Perkin-Elmer, Warrington, Lancashire, U.K.) using M13FOR (SEQ ID No 26) (5'-GTAAAACGACGGCCAGT) and M13REV (SEQ ID No 27) (5'-GGATAACAATTTCACACAGG) oligonucleotides. The pDM12 PDCP display vector DNA sequence between the HindIII and EcoRI restriction sites is shown in Figure 1.

Example 2. Insertion of a random-primed human

lymphocyte cDNA into pDM12 and preparation of a master PDCP stock.

Libraries of peptides can be constructed by many methods known to those skilled in the art. The example given describes a method for constructing a peptide library from randomly primed cDNA, prepared from mRNA isolated from a partially purified cell population.

mRNA was isolated from approximately 10^9 human peripheral blood lymphocytes using a polyATract 1000 mRNA isolation kit (Promega, Southampton, UK). The cell pellet was resuspended in 4ml extraction buffer (4M guanidine thiocyanate, 25mM sodium citrate pH 7.1, 2% β-mercaptoethanol). 8ml of pre-heated (70°C) dilution buffer (6xSSC, 10mM Tris pH 7.4, 1mM EDTA, 0.25% SDS, 1% β-mercaptoethanol) was added to the homogenate and mixed thoroughly by inversion. 10μl of biotinylated oligo-dT (50 pmol/μl) was added, mixed and the mixture incubated at 70°C for 5 minutes. The lymphocyte cell lysate was transferred to 6x 2ml sterile tubes and spun at 13,000 rpm in a microcentrifuge for ten minutes at ambient temperature to produce a cleared lysate. During this centrifugation, streptavidin coated magnetic beads were resuspended and 6ml transferred to a sterile 50ml Falcon tube, then placed in the magnetic stand in a horizontal position until all the beads were captured. The supernatant was carefully poured off and beads resuspended in 6ml 0.5xSSC, then the capture repeated. This wash was repeated 3 times, and beads resuspended in a final volume of 6ml 0.5xSSC. The cleared lysate was added to the washed beads, mixed by inversion and incubated at ambient temperature for 2 minutes, then beads captured in the magnetic stand in a horizontal position. The beads were resuspended gently in 2ml 0.5xSSC and transferred to a sterile 2ml screwtop tube,

then captured again in the vertical position, and the wash solution discarded. This wash was repeated twice more. 1ml of DEPC-treated water was added to the beads and mixed gently. The beads were again captured and the eluted mRNA transferred to a sterile tube. 50μl was electrophoresed to check the quality and quantity of mRNA, while the remainder was precipitated with 0.1 volumes 3M sodium acetate and three volumes absolute ethanol at -80°C overnight in 4 aliquots in sterile 1.5ml screwtop tubes.

Double stranded cDNA was synthesised as described in Example 1 using 5μg of lymphocyte mRNA as template. The cDNA was PCR amplified using oligonucleotides CDNAPCRFOR (SEQ ID No 28) (5'-AAAGCGGCCGCACTGGCCTGAGAGA), which anneals to the cDNA synthesis oligonucleotide described in Example 1 which is present at the 3'-end of all synthesised cDNA molecules incorporates a NotI restriction site, and an equimolar mixture of CDNAPCRBAK1, CDNAPCRBAK2 and CDNAPCRBAK3.
CDNAPCRBAK1: (SEQ ID No 29) 5'-AAAAGGCCCAGCCGGCCATGGCCCAGCCCACCACGCGTCCG,
CDNAPCRBAK2: (SEQ ID No 30) 5'-AAAAGGCCCAGCCGGCCATGGCCCAGTCCCACCACGCGTCCG,
CDNAPCRBAK3: (SEQ ID No 31) 5'-AAAAGGCCCAGCCGGCCATGGCCCAGTACCCACCACGCGTCCG),
all three of which anneal to the SalI adaptor sequence found at the 5'-end of the cDNA and incorporate a SfiI restriction site at the cDNA 5'-end. Ten PCR reactions were carried out using 2μl of cDNA (50ng) per reaction as described in Example 1 using 25 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 2 minutes. The reactions were pooled and a 20μl aliquot checked by agarose gel electrophoresis, the remainder was phenol/chloroform extracted and ethanol precipitated and resuspended in

100μl sterile water. 5μg of pDM12 vector DNA and lymphocyte cDNA PCR product were SfiI-NotI digested phenol/chloroform extracted and small DNA fragments removed by size selection on Chromaspin 1000 spin columns (Clontech, Palo Alto, California, U.S.A.) by centrifugation at 700g for 2 minutes at room temperature. Digested pDM12 and lymphocyte cDNA were ethanol precipitated and ligated together for 16 hours at 16°C. The ligated DNA was precipitated and electroporated in to TG1 *E. coli*. Cells were grown in 1ml SOC medium per cuvette used for 1 hour at 37°C, and plated onto 2xTY agar plates supplemented with 1% glucose and 100μg/ml ampicillin. 10^{-4}, 10^{-5} and 10^{-6} dilutions of the electroporated bacteria were also plated to assess library size. Colonies were allowed to grow overnight at 30°C. $2x10^{8}$ ampicillin resistant colonies were recovered on the agar plates.

The bacteria were then scraped off the plates into 40ml 2xTY broth supplemented with 20% glycerol, 1% glucose and 100μg/ml ampicillin. 5ml was added to a 20ml 2xTY culture broth supplemented with 1% glucose and 100μg/ml ampicillin and infected with 10^{11} kanamycin resistance units (kru) M13K07 helper phage at 37°C for 30 minutes without shaking, then for 30 minutes with shaking at 200rpm. Infected bacteria were transferred to 200ml 2xTY broth supplemented with 25μg/ml kanamycin, 100μg/ml ampicillin, and 20μM IPTG, then incubated overnight at 37°C, shaking at 200rpm. Bacteria were pelleted at 4000rpm for 20 minutes in 50ml Falcon tubes, and 40ml 2.5M NaCl/20% PEG 6000 was added to 200ml of particle supernatant, mixed vigorously and incubated on ice for 1 hour to precipitate PDCP particles. Particles were pelleted at 11000rpm for 30 minutes in 250ml Oakridge tubes at 4°C in a Sorvall RC5B centrifuge, then resuspended in 2ml PBS buffer after removing all traces of PEG/NaCl with a pipette,

then bacterial debris removed by a 5 minute 13500rpm spin in a microcentrifuge. The supernatent was filtered through a 0.45μm polysulfone syringe filter and stored at -20°C.

Example 3. Isolation of human immunoglobulin kappa light chains by repeated rounds of selection against anti-human kappa antibody.

For the first round of library selection a 70x11mm NUNC Maxisorp Immunotube (Life Technologies, Paisley, Scotland U.K.) was coated with 2.5ml of 10μg/ml of anti-human kappa antibody (Seralab, Crawley Down, Sussex, U.K.) in PBS for 2 hours at 37°C. The tube was rinsed three times with PBS (fill & empty) and blocked with 3ml PBS/2% BSA for 2 hours at 37°C and washed as before. $4x10^{12}$ a.r.u. of pDM12-lymphocyte cDNA PDCP stock was added in 2ml 2% BSA/PBS/0.05% Tween 20, and incubated for 30 minutes on a blood mixer, then for 90 minutes standing at ambient temperature. The tube was washed ten times with PBS/0.1% Tween 20, then a further ten times with PBS only. Bound particles were eluted in 1ml of freshly prepared 0.1M triethylamine for 10 minutes at ambient temperature on a blood mixer. Eluted particles were transferred to 0.5ml 1M Tris pH 7.4, vortex mixed briefly and transferred to ice.

Neutralised particles were added to 10ml log phase TG1 E coli bacteria (optical density: OD_{600nm} 0.3-0.5) and incubated at 37°C without shaking for 30 minutes, then with shaking at 200rpm for 30 minutes. 10^{-3}, 10^{-4} & 10^{-5} dilutions of the infected culture were prepared to estimate the number of particles recovered, and the remainder was spun at 4000 rpm for 10 minutes, and the pellet resuspended in 300μl 2xTY medium by vortex mixing. Bacteria were plated onto 2xTY agar plates

supplemented with 1% glucose and $100\mu g/ml$ ampicillin. Colonies were allowed to grow overnight at 30°C.

A PDCP stock was prepared from the bacteria recovered from the first round of selection, as described in Example 2 from a 100ml overnight culture. $250\mu l$ of the round 1 amplified PDCP stock was then selected against anti-human kappa antibody as described above with the tube was washed twelve times with PBS/0.1% Tween 20, then a further twelve times with PBS only.

To identify selected clones, eighty-eight individual clones recovered from the second round of selection were then tested by ELISA for binding to anti-human kappa antibody. Individual colonies were picked into $100\mu l$ 2xTY supplemented with $100\mu g/ml$ ampicillin and 1% glucose in 96-well plates (Costar) and incubated at 37°C and shaken at 200rpm for 4 hours. $25\mu l$ of each culture was transferred to a fresh 96-well plate, containing $25\mu l$/well of the same medium plus 10^7 k.r.u. M13K07 kanamycin resistant helper phage and incubated at 37°C for 30 minutes without shaking, then incubated at 37°C and shaken at 200rpm for a further 30 minutes. $160\mu l$ of 2xTY supplemented with $100\mu g/ml$ ampicillin, $25\mu g/ml$ kanamycin, and $20\mu M$ IPTG was added to each well and particle amplification continued for 16 hours at 37°C while shaking at 200rpm. Bacterial cultures were spun in microtitre plate carriers at 2000g for 10 minutes at 4°C in a benchtop centrifuge to pellet bacteria and culture supernatant used for ELISA.

A Dynatech Immulon 4 ELISA plate was coated with 200ng/well anti-human kappa antibody in $100\mu l$ /well PBS for one hour at 37°C. The plate was washed $2x200\mu l$/well PBS and blocked for 1 hour at 37°C with $200\mu l$/well 2% BSA/PBS and then washed $2x200\mu l$/well PBS. $50\mu l$ PDCP

culture supernatant was added to each well containing 50μl/well 4% BSA/PBS/0.1%Tween 20, and allowed to bind for 1 hour at ambient temperature. The plate was washed three times with 200μl/well PBS/0.1% Tween 20, then three times with 200μl/well PBS. Bound PDCPs were detected with 100μl/well, 1:5000 diluted anti-M13-HRP conjugate (Pharmacia) in 2% BSA/PBS/0.05% Tween 20 for 1 hour at ambient temperature and the plate washed six times as above. The plate was developed for 5 minutes at ambient temperature with 100μl/well freshly prepared TMB (3,3',5,5'-Tetramethylbenzidine) substrate buffer (0.005% H_2O_2, 0.1mg/ml TMB in 24mM citric acid/52mM sodium phosphate buffer pH 5.2). The reaction was stopped with 100μl/well 12.5% H_2SO_4 and read at 450nm. (ELISA data for binding clones is shown in Figure 2).

These clones were then sequenced with M13REV primer (SEQ ID No 27) as in Example 1. The sequence of two of the clones isolated is shown in Figure 3 (see SEQ ID Nos 7 to 10).

Example 4: Construction of the pDM14 N-terminal display vector

It would be useful to design vectors that contain a second DBD binding sequence, such as a second oestrogen receptor HRE sequence, thus allowing the display of increased numbers of peptides per PDCP. Peale et al. (1988, Proc. Natl. Acad. Sci. USA 85: 1038-1042) describe a number of oestrogen receptor HRE sequences. These sequences were used to define an HRE sequence, which differs from that cloned in pDM12, which we used to create a second N-terminal display vector (pDM14). The oligonucleotide: 5'-AAAAGAATTCGAGGTTACATTAACTTTGTT CCGGTCAGACTGACCCAAGTCGACCTGAATGTGTTATTTTAG-3' (SEQ ID No 32) was synthesised and used to mutagenise pDM12 by

PCR with pDM12BAK oligonucleotide as described in Example 1 using 100ng pDM12 vector DNA as template. The resulting DNA fragment, which contained the oestrogen receptor DBD and two HRE sequences separated by a SalI restriction enzyme site, was NotI-EcoRI restriction enzyme digested and cloned into NotI-EcoRI digested pDM12 vector DNA as described in Example 1 to create pDM14. The sequence of pDM14 between the HindIII and EcoRI restriction enzyme sites was checked by DNA sequencing. The final vector sequence between these two sites is shown in Figure 4 (see SEQ ID Nos 11 and 12).

Example 5. Construction of the pDM16 C-terminal display vector

In order to demonstrate the display of peptides fused to the C-terminus of a DBD on a PDCP a suitable vector, pDM16, was created.

In pDM16 the pelB leader DNA sequence is fused directly to the oestrogen receptor DBD sequence removing the multiple cloning sites and the Gly_4Ser linker DNA sequence found in pDM12 and pDM14, which are appended to the C-terminal end of the DBD sequence upstream of the HRE DNA sequence.

To create this vector two separate PCR reactions were carried out on a Techne Progene thermal cycler for 30 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 1 minute. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel using a Mermaid or Geneclean II kit, respectively, according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

In the first, the 5'-untranslated region and pelB

leader DNA sequence was amplified from 100ng of pDM12 vector DNA using 50pmol of each of the oligonucleotides pelBFOR (SEQ ID No 33) (5'-CCTTGGCAGATTCCATCT CGGCCATTGCCGGC-3') and M13REV (SEQ ID NO 27) (see above) in a 100μl reaction containing 0.1mM dNTPs, 2.5 units Taqplus DNA polymerase, and 1x High Salt PCR reaction buffer (20mM Tris-HCl pH 9.2, 60mM KCl, 2mM $MgCl_2$) (Stratagene Ltd, Cambridge, U.K.).

In the second, the 3'-end of the pelB leader sequence and the oestrogen receptor DBD was amplified from 100ng of pDM12 vector DNA using 50pmol of each of the oligonucleotides pelBBAK (SEQ ID No 34) (5'-CCGGCAA TGGCCGAGATGGAATCTGCCAAGG-3') and pDM16FOR (SEQ ID No 35) (5'-TTTTGTCGACTCAATCAGTTATGCGGCCGCCAGCTGCAGG AGGGCCGGCTGGGCCGACCCTCCTCCCCAGACCCCACTTCACCCC-3') in a 100μl reaction containing 0.1mM dNTPs, 2.5 units Taqplus DNA polymerase, and 1x High Salt PCR reaction buffer (Stratagene Ltd, Cambridge, U.K.). Following gel purification both products were mixed together and a final round of PCR amplification carried out to link the two products together as described above, in a 100μl reaction containing 0.1mM dNTPs, 2.5 units Taq DNA polymerase, and 1x PCR reaction buffer (10mM Tris-HCl pH 9.0, 5mM KCl, 0.01% Triton X*-100, 1.5mM $MgCl_2$) (Promega Ltd, Southampton, U.K.).

The resulting DNA fragment, was HindIII-SalI restriction enzyme digested and cloned into HindIII-SalI digested pDM14 vector DNA as described in Example 1 to create pDM16. The sequence of pDM16 between the HindIII and EcoRI restriction enzyme sites was checked by DNA sequencing. The final vector sequence between these two sites is shown in Figure 5 (see SEQ ID Nos 13 and 14).

Example 6. Display of the C-terminal fragment of human N-cadherin on the surface of a PDCP

cDNA libraries of peptides can be constructed by many methods known to those skilled in the art. One commonly used method for constructing a peptide library uses oligo dT primed cDNA, prepared from polyA+ mRNA. In this method the first-strand synthesis is carried out using an oligonucleotide which anneals to the 3'-end polyA tail of the mRNA composed of T_n (where n is normally between 10 and 20 bases) and a restriction enzyme site such as NotI to facilitate cloning of cDNA. The cDNA cloned by this method is normally composed of the polyA tail, the 3'- end untranslated region and the C-terminal coding region of the protein. As an example of the C-terminal display of peptides on a PDCP, a human cDNA isolated from a library constructed by the above method was chosen.

The protein N-cadherin is a cell surface molecule involved in cell-cell adhesion. The C-terminal cytoplasmic domain of the human protein (Genbank database accession number: M34064) is recognised by a commercially available monoclonal antibody which was raised against the C-terminal 23 amino acids of chicken N-cadherin (SIGMA catalogue number: C-1821). The 1.4kb human cDNA fragment encoding the C-terminal 99 amino acids, 3'- untranslated region and polyA tail (NotI site present at the 3'-end of the polyA tail) was amplified from approximately 20ng pDM7-NCAD#C with 25pmol of each oligonucleotide M13FOR (SEQ ID No 26) and CDNPCRBAK1 (SEQ ID No 29) (see above) in a 50μl reaction containing 0.1mM dNTPs, 2.5 units Taqplus DNA polymerase, and 1x High Salt PCR reaction buffer (20mM Tris-HCl pH 9.2, 60mM KCl, 2mM $MgCl_2$) (Stratagene Ltd, Cambridge, U.K.) on a Techne Progene thermal cycler for

30 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 1 minute. Following gel purification and digestion with SfiI and NotI restiction enzymes, the PCR product was cloned into pDM16 using an analogous protocol as described in Example 1.

Clones containing inserts were identified by ELISA of 96 individual PDCP cultures prepared as described in Example 3. A Dynatech Immulon 4 ELISA plate was coated with 1:250 diluted anti-pan cadherin monoclonal antibody in 100μl /well PBS overnight at 4ºC. The plate was washed 3x200μl/well PBS and blocked for 1 hour at 37°C with 200μl/well 2% Marvel non-fat milk powder/PBS and then washed 2x200μl/well PBS. 50μl PDCP culture supernatant was added to each well containing 50μl/well 4% Marvel/PBS, and allowed to bind for 1 hour at ambient temperature. The plate was washed three times with 200μl/well PBS/0.1% Tween 20, then three times with 200μl/well PBS. Bound PDCPs were detected with 100μl/well, 1:5000 diluted anti-M13-HRP conjugate (Pharmacia) in 2% Marvel/PBS for 1 hour at ambient temperature and the plate washed six times as above. The plate was developed for 15 minutes at ambient temperature with 100μl/well freshly prepared TMB (3,3',5,5'-Tetramethylbenzidine) substrate buffer (0.005% H_2O_2, 0.1mg/ml TMB in 24mM citric acid/52mM sodium phosphate buffer pH 5.2). The reaction was stopped with 100μl/well 12.5% H_2SO_4 and read at 450nm. The nucleotide sequence of an ELISA positive clone insert and DBD junction was checked by DNA sequencing using oligonucleotides M13FOR (SEQ ID No 26) (see Example 1) and ORSEQBAK (SEQ ID No 36) (5'-TGTTGAAACACAAGCGCCAG-3').

A fifty-fold concentrated stock of C-terminal N-cadherin PDCP particles was prepared by growing the un-

infected TG1 clone in 1ml 2xTY culture broth supplemented with 1% glucose and 100μg/ml ampicillin for five hours at 37°C, shaking at 200rpm and infecting with 10^8 kanamycin resistance units (kru) M13K07 helper phage at 37°C for 30 minutes without shaking, then for 30 minutes with shaking at 200rpm. Infected bacteria were transferred to 20ml 2xTY broth supplemented with 25μg/ml kanamycin, 100μg/ml ampicillin, and 20μM IPTG, then incubated overnight at 30°C, shaking at 200rpm. Bacteria were pelleted at 4000rpm for 20 minutes in 50ml Falcon tubes, and 4ml 2.5M NaCl/20% PEG 6000 was added to 20ml of PDCP supernatant, mixed vigorously and incubated on ice for 1 hour to precipitate particles.

The particles were pelleted at 11000rpm for 30 minutes in 50ml Oakridge tubes at 4°C in a Sorvall RC5B centrifuge, then resuspended in PBS buffer after removing all traces of PEG/NaCl with a pipette, then bacterial debris removed by a 5 minute 13500rpm spin in a microcentrifuge. The supernatant was filtered through a 0.45μm polysulfone syringe filter. The concentrated stock was two-fold serially diluted and used in ELISA against plates coated with anti-pan-cadherin antibody as described above (see Figure 6).

This example demonstrates the principle of C-terminal display using PDCPs, that C-terminal DBD-peptide fusion PDCPs can be made which can be detected in ELISA, and the possibility that oligo dT primed cDNA libraries may be displayed using this method.

Example 7. Display of *in vivo* biotinylated C-terminal domain of human propionyl CoA carboxylase on the surface of a PDCP

Example 6 shows that the C-terminal domain of human N-

cadherin can be expressed on the surface of a PDCP as a C-terminal fusion with the DBD. Here it is shown that the C-terminal domain of another human protein propionyl CoA carboxylase alpha chain (Genbank accession number: X14608) can similarly be displayed, suggesting that this methodology may be general.

The alpha sub-unit of propionyl CoA carboxylase alpha chain (PCC) contains 703 amino acids and is normally biotinylated at position 669. It is demonstrated that the PCC peptide displayed on the PDCP is biotinylated, as has been shown to occur when the protein is expressed in bacterial cells (Leon-Del-Rio & Gravel; 1994, J. Biol. Chem. 37, 22964-22968).

The 0.8kb human cDNA fragment of PCC alpha encoding the C-terminal 95 amino acids, 3'- untranslated region and polyA tail (NotI site present at the 3'-end of the polyA tail) was amplified and cloned into pDM16 from approximately 20ng pDM7-PCC#C with 25pmol of each oligonucleotide M13FOR (SEQ ID No 26) and CDNPCRBAK1 (SEQ ID No 29) as described in Example 6.

Clones containing inserts were identified by ELISA as described in Example 6, except that streptavidin was coated onto the ELISA plate at 250ng/well, in place of the anti-cadherin antibody. The nucleotide sequence of an ELISA positive clone insert and DBD junction was checked by DNA sequencing using oligonucleotides M13FOR (SEQ ID No 26) and ORSEQBAK (SEQ ID No 36) (see above). A fifty-fold concentrated stock of C-terminal PCC PDCP particles was prepared and tested in ELISA against streptavidin as described in Example 6 (see Figure 7).

This example shows not only that the peptide can be displayed as a C-terminal fusion on a PDCP, but also

that *in vivo* modified peptides can be displayed.

Example 8. Construction of a human scFv PDCP display library

This example describes the generation of a human antibody library of scFvs made from an un-immunised human. The overall strategy for the PCR assembly of scFv fragments is similar to that employed by Marks, J. D. et al. 1991, J. Mol. Biol. 222: 581-597. The antibody gene oligonucleotides used to construct the library are derived from the Marke et al., paper and from sequence data extracted from the Kabat database (Kabat, E. A. et al., Sequences of Proteins of Immunological Interest. 4th edition. U.S. Department of Health and Human Services. 1987). The three linker oligonucleotides are described by Zhou et al. (1994, Nucleic Acids Res., 22: 888-889), all oligonucleotides used are detailed in Table 1.

First, mRNA was isolated from peripheral blood lymphocytes and cDNA prepared for four repertoires of antibody genes IgD, IgM, Igκ and Igλ, using four separate cDNA synthesis primers. VH genes were amplified from IgD and IgM primed cDNA, and VL genes were amplified from Igκ and Igλ primed cDNA. A portion of each set of amplified heavy chain or light chain DNA was then spliced with a separate piece of linker DNA encoding the 15 amino acids $(Gly_4\ Ser)_3$ (Huston, J. S. et al. 1989, Gene, 77: 61). The 3'-end of the VH PCR products and the 5'-end of the VL PCR products overlap the linker sequence as a result of incorporating linker sequence in the JH, Vκ and Vλ family primer sets (Table 1). Each VH-linker or linker-VL DNA product was then spliced with either VH or VL DNA to produce the primary scFv product in a VH-linker-VL configuration. This scFv

product was then amplified and cloned into pDM12 as a SfiI-NotI fragment, electroporated into TG1 and a concentrated PDCP stock prepared.

mRNA isolation and cDNA synthesis.

Human lymphocyte mRNA was purified as described in Example 2. Separate cDNA reactions were performed with IGDCDNAFOR (SEQ ID No 37), IGMCDNAFOR (SEQ ID No 38), IGKCDNAFOR (SEQ ID No 39) and IGλCDNAFOR (SEQ ID No 40) oligonucleotides. 50pmol of each primer was added to approximately 5μg of mRNA in 20μl of nuclease free water and heated to 70°C for 5 minutes and cooled rapidly on ice, then made up to a final reaction volume of 100μl containing 50mM Tris pH 8.3, 75mM KCl, 3mM $MgCl_2$, 10mM DTT, 0.5mM dNTPs, and 2000 units of Superscript II reverse transcriptase (Life Technologies, Paisley, Scotland, U.K.). The reactions were incubated at 37°C for two hours, then heated to 95°C for 5 minutes.

Primary PCRs.

For the primary PCR amplifications separate amplifications were set up for each family specific primer with either an equimolar mixture of the JHFOR primer set (SEQ ID Nos 41 to 44) for IgM and IgD cDNA, or with SCFVκFOR (SEQ ID No 51) or SCFVλFOR (SEQ ID No 52) for IgK or Igλ cDNA respectively e.g. VH1BAK and JHFOR set; Vκ2BAK (SEQ ID No 54) and SCFVκFOR (SEQ ID No 51); Vλ3aBAK (SEQ ID No 66) and SCFVλFOR (SEQ ID No 52) etc. Thus, for IgM, IgD and Igκ cDNA six separate reactions were set up, and seven for Igλ cDNA. A 50μl reaction mixture was prepared containing 2μl cDNA, 25pmol of the appropriate FOR and BAK primers, 0.1mM dNTPs, 2.5 units Taqplus DNA polymerase, and 1x High Salt PCR reaction buffer (20mM Tris-HCl pH 9.2, 60mM KCl, 2mM $MgCl_2$) (Stratagene Ltd, Cambridge, U.K.).

Reactions were amplified on a Techne Progene thermal cycler for 30 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 2 minutes, followed by 10 minutes at 72°C. Fifty microlitres of all 25 reaction products were electrophoresed on an agarose gel, excised and products purified from the gel using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.). All sets of IgD, IgM, IgK or Igλ reaction products were pooled to produce VH or VL DNA sets for each of the four repertoires. These were then adjusted to approximately 20ng/μl.

Preparation of linker.

Linker product was prepared from eight 100μl reactions containing 5ng LINKAMP3T (SEQ ID No 76) template oligonucleotide, 50pmol of LINKAMP3 (SEQ ID No 74) and LINKAMP5 (SEQ ID No 75) primers, 0.1mM dNTPs, 2.5 units Taqplus DNA polymerase, and 1x High Salt PCR reaction buffer (20mM Tris-HCl pH 9.2, 60mM KCl, 2mM $MgCl_2$) (Stratagene Ltd, Cambridge, U.K.). Reactions were amplified on a Techne Progene thermal cycler for 30 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 1 minute, followed by 10 minutes at 72°C. All reaction product was electrophoresed on a 2% low melting point agarose gel, excised and products purified from the gel using a Mermaid kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.) and adjusted to 5ng/μl.

First stage linking.

Four linking reactions were prepared for each repertoire using 20ng of VH or VL DNA with 5ng of Linker DNA in 100μl reactions containing (for IgM or IgD VH) 50pmol of LINKAMPFOR and VH1-6BAK set, or, 50pmol LINKAMPBAK and either SCFVκFOR (Igκ) or SCFVλFOR (Igλ), 0.1mM dNTPs, 2.5 units Taq DNA polymerase, and

1x PCR reaction buffer (10mM Tris-HCl pH 9.0, 5mM KCl, 0.01% Triton X®-100, 1.5mM $MgCl_2$) (Promega Ltd, Southampton, U.K.). Reactions were amplified on a Techne Progene thermal cycler for 30 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 2 minutes, followed by 10 minutes at 72°C. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.) and adjusted to 20ng/μl.

Final linking and reamplification.

To prepare the final scFv DNA products, five 100μl reactions were performed for VH-LINKER plus VL DNA, and, five 100μl reactions were performed for VH plus LINKER-VL DNA for each of the four final repertoires (IgM VH-VK, VH-Vλ; IgD VH-VK, VH-Vλ) as described in step (d) above using 20ng of each component DNA as template. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.) and adjusted to 20ng/μl. Each of the four repertoires was then re-amplified in a 100μl reaction volume containing 2ng of each linked product, with 50pmol VHBAK1-6 (SEQ ID Nos 53 to 58) and either the JKFOR (SEQ ID Nos 66 to 70) or JλFOR (SEQ ID Nos 71 to 73) primer sets, in the presence of 0.1mM dNTPs, 2.5 units Taq DNA polymerase, and 1x PCR reaction buffer (10mM Tris-HCl pH 9.0, 5mM KCl, 0.01% Triton X®-100, 1.5mM $MgCl_2$) (Promega Ltd, Southampton, U.K.). Thirty reactions were performed per repertoire to generate enough DNA for cloning.

Reactions were amplified on a Techne Progene thermal cycler for 25 cycles of 94°C, 1 minute; 65°C, 1 minute; 72°C, 2 minutes, followed by 10 minutes at 72°C. Reaction products were phenol-chloroform extracted,

ethanol precipitated, vacuum dried and re-suspended in 80μl nuclease free water.

Cloning into pDM12.

Each of the four repertoires was SfiI-NotI digested, and electrophoresed on an agarose gel, excised and products purified from the gel using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.). Each of the four repertoires was ligated overnight at 16°C in 140μl with 10μg of SfiI-NotI cut pDM12 prepared as in Example 2, and 12 units of T4 DNA ligase (Life Technologies, Paisley, Scotland, U.K.). After incubation the ligations were adjusted to 200μl with nuclease free water, and DNA precipitated with 1μl 20mg/ml glycogen, 100μl 7.5M ammonium acetate and 900μl ice-cold (-20°C) absolute ethanol, vortex mixed and spun at 13,000rpm for 20 minutes in a microfuge to pellet DNA. The pellets were washed with 500μl ice-cold 70% ethanol by centrifugation at 13,000rpm for 2 minutes, then vacuum dried and re-suspended in 10μl DEPC-treated water. 1μl aliquots of each repertoire was electroporated into 80μl *E. coli* (TG1). Cells were grown in 1ml SOC medium per cuvette used for 1 hour at 37°C, and plated onto 2xTY agar plates supplemented with 1% glucose and 100μg/ml ampicillin. 10^{-4}, 10^{-5} and 10^{-6} dilutions of the electroporated bacteria were also plated to assess library size. Colonies were allowed to grow overnight at 30°C. Cloning into SfiI-NotI digested pDM12 yielded an IgM-κ/λ repertoire of 1.16×10^9 clones, and an IgD-κ/λ repertoire of 1.21×10^9 clones.

Preparation of PDCP stock.

Separate PDCP stocks were prepared for each repertoire library. The bacteria were then scraped off the plates into 30ml 2xTY broth supplemented with 20% glycerol, 1%

glucose and 100μg/ml ampicillin. 3ml was added to a 50ml 2xTY culture broth supplemented with 1% glucose and 100μg/ml ampicillin and infected with 10^{11} kanamycin resistance units (kru) M13K07 helper phage at 37°C for 30 minutes without shaking, then for 30 minutes with shaking at 200rpm. Infected bacteria were transferred to 500ml 2xTY broth supplemented with 25μg/ml kanamycin, 100μg/ml ampicillin, and 20μM IPTG, then incubated overnight at 30°C, shaking at 200rpm. Bacteria were pelleted at 4000rpm for 20 minutes in 50ml Falcon tubes, and 80ml 2.5M NaCl/20% PEG 6000 was added to 400ml of particle supernatant, mixed vigorously and incubated on ice for 1 hour to precipitate PDCP particles. Particles were pelleted at 11000rpm for 30 minutes in 250ml Oakridge tubes at 4°C in a Sorvall RC5B centrifuge, then resuspended in 40ml water and 8ml 2.5M NaCl/20% PEG 6000 added to reprecipitate particles, then incubated on ice for 20 minutes. Particles were again pelleted at 11000rpm for 30 minutes in 50ml Oakridge tubes at 4°C in a Sorvall RC5B centrifuge, then resuspended in 5ml PBS buffer, after removing all traces of PEG/NaCl with a pipette. Bacterial debris was removed by a 5 minute 13500rpm spin in a microcentrifuge. The supernatant was filtered through a 0.45μm polysulfone syringe filter, adjusted to 20% glycerol and stored at -70°C.

Example 9. Isolation of binding activity from a N-terminal display PDCP library of human scFvs

The ability to select binding activities to a target of interest from a human antibody library is important due to the possibility of generating therapeutic human antibodies. In addition, such libraries allow the isolation of antibodies to targets which cannot be used for traditional methods of antibody generation due to

toxicity, low immunogenicity or ethical considerations. In this example we demonstrate the isolation of specific binding activities against a peptide antigen from a PDCP library of scFvs from an un-immunised human.

The generation of the library, used for the isolation of binding activities in this example, is described in Example 8.

Substance P is an eleven amino acid neuropeptide involved in inflammatory and pain responses *in vivo*. It has also been implicated in a variety of disorders such as psoriasis and asthma amongst others (Misery, L. 1997, Br. J. Dertmatol., 137: 843-850; Maggi, C. A. 1997, Regul. Pept. 70: 75-90; Choi, D. C. & Kwon, O.J., 1998, Curr. Opin. Pulm. Med., 4: 16-24). Human antibodies which neutralise this peptide may therefore have some therapeutic potential. As this peptide is too small to coat efficiently on a tube, as described in Example 3, selection of binding activities was performed in-solution, using N-terminal biotinylated substance P and capturing bound PDCP particles on streptavidin-coated magnetic beads.

Enrichment for substance P binding PDCP particles.
An aliquot of approximately 10^{13} a.r.u. IgM and IgD scFv library stock was mixed with $1\mu g$ biotinylated substance P in $800\mu l$ 4% BSA/0.1% Tween 20/PBS, and allowed to bind for two hours at ambient temperature. Bound PDCPs were then captured onto 1ml of BSA blocked streptavidin coated magnetic beads for 10 minutes at ambient temperature. The beads were captured to the side of the tube with a magnet (Promega), and unbound material discarded. The beads were washed eight times with 1ml PBS/0.1% Tween 20/ $10\mu g/ml$ streptavidin, then two times

with 1ml of PBS by magnetic capture and removal of wash buffer. After the final wash bound PDCPs were eluted with 1ml of freshly prepared 0.1M triethylamine for 10 minutes, the beads were captured, and eluted particles transferred to 0.5ml 1M Tris-HCl pH 7.4. Neutralised particles were added to 10ml log phase TG1 *E. coli* bacteria and incubated at 37°C without shaking for 30 minutes, then with shaking at 200rpm for 30 minutes. 10^{-3}, 10^{-4} & 10^{-5} dilutions of the infected culture were prepared to estimate the number of particles recovered, and the remainder was spun at 4000 rpm for 10 minutes, and the pellet resuspended in 300μl 2xTY medium by vortex mixing. Bacteria were plated onto 2xTY agar plates supplemented with 1% glucose and 100μg/ml ampicillin. Colonies were allowed to grow overnight at 30°C. A 100-fold concentrated PDCP stock was prepared from a 200ml amplified culture of these bacteria as described above, and 0.5ml used in as second round of selection with 500ng biotinylated substance P. For this round 100μg/ml streptavidin was included in the wash buffer.

ELISA identification of binding clones.

Binding clones were identified by ELISA of 96 individual PDCP cultures prepared as described in Example 3 from colonies recovered after the second round of selection. A Dynatech Immulon 4 ELISA plate was coated with 200ng/well streptavidin in 100μl /well PBS for 1 hour at 37°C. The plate was washed 3x200μl/well PBS and incubated with 10ng/well biotinylated substance P in 100μl /well PBS for 30 minutes at 37°C The plate was washed 3x200μl/well PBS and blocked for 1 hour at 37°C with 200μl/well 2% Marvel non-fat milk powder/PBS and then washed 2x200μl/well PBS. 50μl PDCP culture supernatant was added to each well containing 50μl/well 4% Marvel/PBS,

and allowed to bind for 1 hour at ambient temperature. The plate was washed three times with 200μl/well PBS/0.1% Tween 20, then three times with 200μl/well PBS. Bound PDCPs were detected with 100μl/well, 1:5000 diluted anti-M13-HRP conjugate (Pharmacia) in 2% Marvel/PBS for 1 hour at ambient temperature and the plate washed six times as above. The plate was developed for 10 minutes at ambient temperature with 100μl/well freshly prepared TMB (3,3',5,5'-Tetramethylbenzidine) substrate buffer (0.005% H_2O_2, 0.1mg/ml TMB in 24mM citric acid/52mM sodium phosphate buffer pH 5.2). The reaction was stopped with 100μl/well 12.5% H_2SO_4 and read at 450nm. Out of 96 clones tested, 10 gave signals greater than twice background (background = 0.05).

Characterization of a binding clone.
A 50-fold concentrated PDCP stock was prepared from a 100ml amplified culture of a single ELISA positive clone as described above. 10μl per well of this stock was tested in ELISA as described above for binding to streptavidin, streptavidin-biotinylated-substance P and streptavidin-biotinylated-CGRP (N-terminal biotinylated). Binding was only observed in streptavidin-biotinylated-substance P coated wells indicating that binding was specific. In addition, binding to streptavidin-biotinylated substance P was completely inhibited by incubating the PDCP with 1μg/ml free substance P (see Figure 8). The scFv VH (SEQ ID Nos 15 and 16) and VL (SEQ ID Nos 17 and 18) DNA and amino acid sequence was determined by DNA sequencing with oligonucleotides M13REV (SEQ ID No27) and ORSEQFOR (SEQ ID No 36) and is shown in Figure 9.

The results indicate that target binding activities can be isolated from PDCP display libraries of human scFv

fragments.

Example 10

In another example the invention provides methods for screening a DNA library whose members require more than one chain for activity, as required by, for example, antibody Fab fragments for ligand binding. To increase the affinity of an antibody of known heavy and light chain sequence, libraries of unknown light chains co-expressed with a known heavy chain are screened for higher affinity antibodies. The known heavy chain antibody DNA sequence is joined to a nucleotide sequence encoding a oestrogen receptor DNA binding domain in a phage vector which does not contain the oestrogen receptor HRE sequence. The antibody DNA sequence for the known heavy chain (VH and CH1) gene is inserted in the 5′ region of the oestrogen receptor DBD DNA, behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of the downstream fusion protein to the periplasmic space. The library of unknown light chains (VL and CL) is expressed separately from a phagemid expression vector which also contains the oestrogen receptor HRE sequence. Thus when both heavy and light chains are expressed in the same host cell, following infection with the phage containing the heavy chain-DBD fusion, the light chain phagemid vector is preferentially packaged into mature phage particles as single stranded DNA, which is bound by the heavy chain-DBD fusion protein during the packaging process. The light chain proteins are transported to the periplasm where they assemble with the heavy chain that is fused to the DBD protein as it exits the cell on the PDCP. In this example the DBD fusion protein and the HRE DNA sequences are not encoded on the same vector, the unknown peptide

sequences are present on the same vector as the HRE sequence. Peptide display carrier packages (PDCP) which encode the protein of interest can then be selected by means of a ligand specific for the antibody.

Table 1 (i) Oligonucleotide primers used for human scFv library construction

cDNA synthesis primers

IgMCDNAFOR	TGGAAGAGGCACGTTCTTTTCTTT
IgDCDNAFOR	CTCCTTCTTACTCTTGCTGGCGGT
IgκCDNAFOR	AGACTCTCCCCTGTTGAAGCTCTT
IgλCDNAFOR	TGAAGATTCTGTAGGGGCCACTGTCTT

JHFOR primers

JH1-2FOR	TGAACCGCCTCCACCTGAGGAGACGGTGACCAGGGTGCC
JH3FOR	TGAACCGCCTCCACCTGAAGAGACGGTGACCATTGTCCC
JH4-5FOR	TGAACCGCCTCCACCTGAGGAGACGGTGACCAGGGTTCC
JH6FOR	TGAACCGCCTCCACCTGAGGAGACGGTGACCGTGGTCCC

VH familyBAKprimers

VH1BAK	TTTTTGGCCCAGCCGGCCATGGCCCAGGTGCAGCTGGTGCAGTCTGG
VH2BAK	TTTTTGGCCCAGCCGGCCATGGCCCAGGTCAACTTAAGGGAGTCTGG
VH3BAK	TTTTTGGCCCAGCCGGCCATGGCCGAGGTGCAGCTGGTGGAGTCTGG
VH4BAK	TTTTTGGCCCAGCCGGCCATGGCCCAGGTGCAGCTGCAGGAGTCGGG
VH5BAK	TTTTTGGCCCAGCCGGCCATGGCCGAGGTGCAGCTGTTGCAGTCTGC
VH6BAK	TTTTTGGCCCAGCCGGCCATGGCCCAGGTACAGCTGCAGCAGTCAGG

Light chain FOR primers

SCFVKFOR	TTATTCGCGGCCGCCTAAACAGAGGCAGTTCCAGATTTC
SCFVλFOR	GTCACTTGCGGCCGCCTACAGTGTGGCCTTGTTGGCTTG

VK family BAK primers

VK1BAK	TCTGGCGGTGGCGGATCGGACATCCAGATGACCCAGTCTCC
VK2BAK	TCTGGCGGTGGCGGATCGGATGTTGTGATGACTCAGTCTCC
VK3BAK	TCTGGCGGTGGCGGATCGGAAATTGTGTTGACGCAGTCTCC
VK4BAK	TCTGGCGGTGGCGGATCGGACATCGTGATGACCCAGTCTCC
VK5BAK	TCTGGCGGTGGCGGATCGGAAACGACACTCACGCAGTCTCC
VK6BAK	TCTGGCGGTGGCGGATCGGAAATTGTGCTGACTCAGTCTCC

JK FOR primers

JK1FOR	TTCTCGTGCGGCCGCCTAACGTTTGATTTCCACCTTGGTCCC
JK2FOR	TTCTCGTGCGGCCGCCTAACGTTTGATCTCCAGCTTGGTCCC
JK3FOR	TTCTCGTGCGGCCGCCTAACGTTTGATATCCACTTTGGTCCC
JK4FOR	TTCTCGTGCGGCCGCCTAACGTTTGATCTCCACCTTGGTCCC
JK5FOR	TTCTCGTGCGGCCGCCTAACGTTTAATCTCCAGTCGTGTCCC

Vλ family BAK primers

Vλ1BAK	TCTGGCGGTGGCGGATCGCAGTCTGTGTTGACGCAGCCGCC
Vλ2BAK	TCTGGCGGTGGCGGATCGCAGTCTGCCCTGACTCAGCCTGC

Table 1 (ii) Oligonucleotide primers used for human scFv library construction

Vλ3aBAK	TCTGGCGGTGGCGGATCGTCCTATGTGCTGACTCAGCCACC
Vλ3bBAK	TCTGGCGGTGGCGGATCGTCTTCTGAGCTGACTCAGGACCC
Vλ4BAK	TCTGGCGGTGGCGGATCGCACGTTATACTGACTCAACCGCC
Vλ5BAK	TCTGGCGGTGGCGGATCGCAGGCTGTGCTCACTCAGCCGTC
Vλ6BAK	TCTGGCGGTGGCGGATCGAATTTTATGCTGACTCAGCCCCA

Jλ primers

Jλ1FOR	TTCTCGTGCGGCCGCCTAACCTAGGACGGTGACCTTGGTCCC
Jλ2-3FOR	TTCTCGTGCGGCCGCCTAACCTAGGACGGTCAGCTTGGTCCC
Jλ4-5FOR	TTCTCGTGCGGCCGCCTAACCTAAAACGGTGAGCTGGGTCCC

Linker primers

LINKAMP3	CGATCCGCCACCGCCAGA
LINKAMP5	GTCTCCTCAGGTGGAGGC
LINKAMP3T	CGATCCGCCACCGCCAGAGCCACCTCCGCCTGAACCGCCTCCACCTGAGGAGAC